


An optimized FAIRE procedure for low cell numbers in yeast

David Segorbe¹ | Derek Wilkinson¹ | Alexandru Mizeranschi¹ | Timothy Hughes² |
Ragnhild Aaløkken² | Libuše Váchová³ | Zdena Palková¹ | Gregor D. Gilfillan² 

¹Faculty of Science, Charles University, BIOCEV, 252 50 Vestec, Czech Republic

²Department of Medical Genetics, Oslo University Hospital and University of Oslo, 0450 Oslo, Norway

³Institute of Microbiology of the Czech Academy of Sciences, BIOCEV, 252 50 Vestec, Czech Republic

Correspondence

Gregor D. Gilfillan, Department of Medical Genetics, Oslo University Hospital and University of Oslo, 0450 Oslo, Norway.
Email: gregor.gilfillan@medisin.uio.no

Funding information

Norway Grants, Grant/Award Number: 7F14083 under contract no. MSMT-28477/2014

Abstract

We report an optimized low-input FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements-sequencing) procedure to assay chromatin accessibility from limited amounts of yeast cells. We demonstrate that the method performs well on as little as 4 mg of cells scraped directly from a few colonies. Sensitivity, specificity and reproducibility of the scaled-down method are comparable with those of regular, higher input amounts, and allow the use of 100-fold fewer cells than existing procedures. The method enables epigenetic analysis of chromatin structure without the need for cell multiplication of exponentially growing cells in liquid culture, thus opening the possibility of studying colony cell subpopulations, or those that can be isolated directly from environmental samples.

KEYWORDS

chromatin-accessibility, epigenetics, HTS, NGS, *Saccharomyces cerevisiae*

1 | INTRODUCTION

Eukaryotic genomic regions that are accessed by regulatory proteins or the transcription apparatus must be partially released from their chromatin packaging. Accordingly, chromatin accessibility has been recognized as a proxy for actively transcribed genomic regions and regulatory elements (Tsompana & Buck, 2014). A number of genome-wide assays of chromatin accessibility exist, including MNase-seq, DNase-seq, Sono-seq, FAIRE-seq and ATAC-seq (Auerbach et al., 2009; Boyle et al., 2008; Buenrostro, Giresi, Zaba, Chang, & Greenleaf, 2013; Gaulton et al., 2010; Schones et al., 2008). These assays employ physical, chemical or enzymatic methods to release the least tightly packed chromatin fragments (accessible chromatin), or in the case of MNase-seq, digest the accessible chromatin, leaving the remainder. Libraries of released and isolated DNA fragments can then be prepared for high-throughput sequencing (the exception being ATAC-seq, in which case library preparation by transposase insertion is the procedure which releases accessible chromatin; Buenrostro et al., 2013). Mapping

of sequencing reads then allows determination of the location and extent of open chromatin. For an informative comparison of the above methods, readers are directed to the review by Tsompana and Buck (2014).

FAIRE relies on the differential affinity of nucleosomal vs. free DNA for organic solvent (Hogan, Lee, & Lieb, 2006; Nagy, Cleary, Brown, & Lieb, 2003). Specifically, formaldehyde cross-linking is employed to fix *in vivo* protein–DNA interactions, and subsequent sonication employed to fragment and solubilize the chromatin. Histone–DNA interactions are particularly stable, so nucleosomal DNA is cross-linked to histone proteins with greater efficiency than other protein–DNA interactions. Fixed chromatin fragments are then subjected to repeated phenol-chloroform extraction, with nucleosomal DNA partitioning with the organic–aqueous interface. In contrast, free DNA molecules (open chromatin) associate with the aqueous phase and are preferentially recovered.

FAIRE typically requires fewer input cells than the MNase and DNase I methods, and has the added advantage that less optimization

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2018 The Authors. *Yeast* published by John Wiley & Sons, Ltd.

is required (in particular, it avoids enzyme titration; Tsompana & Buck, 2014). However, FAIRE suffers from a lower signal-to-noise ratio than other methods, and has largely been supplanted by ATAC-seq (Buenrostro et al., 2013), which offers an improved signal-to-noise ratio, is faster to perform and requires even fewer cells. Nonetheless, FAIRE retains one key advantage of particular relevance in yeast: ATAC-seq in yeast requires the creation of spheroplasts to permeabilize cells to the Tn5 transposase (Schep et al., 2015). The time and temperatures required to incubate cells for spheroplast creation may allow epigenetic changes to occur, which is not a concern with FAIRE. Of particular relevance, the efficiency of spheroplast formation varies between cell types and ages (Klis, Mol, Hellingwerf, & Brul, 2002). We therefore chose to use the FAIRE assay to study yeast chromatin accessibility.

The FAIRE assay was devised and first tested in yeast cells (Giresi, Kim, McDaniel, Iyer, & Lieb, 2007; Nagy et al., 2003). However, as in the initial study, those performing the technique have continued to use relatively large numbers of cells from liquid cultures (Berchowitz, Hanlon, Lieb, & Copenhaver, 2009; Connelly, Wakefield, & Akey, 2014; Hogan et al., 2006; Simon, Giresi, Davis, & Lieb, 2012). We present here an optimized FAIRE procedure that works on low-input amounts of cells (down to 4 mg of wet biomass, equivalent to a few colonies), removing the need for amplification in liquid culture. We estimate that this represents an ~100-fold reduction in input relative to previously published procedures.

2 | MATERIALS AND METHODS

2.1 | Yeast strain and culture

The *Saccharomyces cerevisiae* laboratory strain BY4742 (Mata, his3 Δ , leu Δ , lys2 Δ , ura Δ) was obtained from Euroscarf (www.euroscarf.de). For FAIRE experiments, 3-day-old colonies grown at 28°C on GMA plates (3% glycerol, 1% yeast extract, 2% agar) were collected in sterile 50/1.5 mL tubes for further processing.

2.2 | FAIRE assay

Experiments were performed following a modified protocol based on Simon, Giresi, Davis, and Lieb (2013). For detailed instructions on performing the procedure, see the protocol provided as Supporting Information file 1. For the experiments reported here, yeast cells (three independent biological replicates each of 500, 100, 20 and 4 mg) were harvested by collecting multiple colonies by scraping with a spatula and fixed in 200 vols (w/v relative to initial weight harvested) of formaldehyde fixing buffer (1% final concentration formaldehyde, 0.05 M HEPES pH 7.4, 0.1 M NaCl, 10 mM EDTA) for 10 min at 28°C with gentle shaking. Cross-linking was quenched by adding glycine to 125 mM and continuing incubation for 5 min at room temperature. Cells were collected by centrifugation for 10 min at 4000 g at room temperature, and cell pellets were washed three times with 10 mL of phosphate-buffered saline, pH 7.4. Cells were then resuspended in 17 vols of lysis buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl, 2% Triton X100 and 1% SDS) relative to original weights of cell pellets, and transferred to 1.5 mL screw-cap

tubes (up to 250 μ L per tube). Chilled acid-washed beads (Sigma Aldrich, Germany) were added and cells were disrupted with a FastPrep FP120 cell disruptor (MP Biomedicals, Santa Ana, CA, USA), using six cycles of 20 s shaking, with a 2 min pause on ice between cycles. Beads were washed with a further 550 μ L lysis buffer, centrifuged for 5 min at 3000 g at 4°C, and supernatant recovered. Supernatants were adjusted to 600 μ L with additional lysis buffer and divided into two 1.5 mL Bioruptor microtubes (Diagenode, Liege, Belgium). Chromatin was sheared in a Bioruptor Pico sonicator (Diagenode), using three cycles of 30 s on and 30 s off at 4°C to achieve fragments of 300–400 BP, then divided into two aliquots for further processing as input or FAIRE samples.

Input DNA samples were de-cross-linked by adding RNase A to 0.1 mg/mL and incubating at 37°C for 30 min, followed by addition of Proteinase K to 0.2 mg/mL with overnight incubation at 65°C. Finally, input DNA samples were purified by three rounds of phenol chloroform extraction following by ethanol precipitation. FAIRE DNA samples were first subjected to three rounds of phenol chloroform extraction, followed by ethanol precipitation. Then, chromatin was de-cross-linked with RNase A and Proteinase K as above. Finally, FAIRE and input DNA samples were purified using CHIP DNA clean and concentrator columns (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions.

2.3 | Sequencing

FAIRE and input DNA samples were prepared for sequencing using ThruPLEX reagents (Rubicon Genomics, Ann Arbor, MI, USA) according to the manufacturer's instructions. An aliquot of 2–5 ng FAIRE or input control DNA was used as starting material for library preparation. Sequencing was performed on a NextSeq 500 instrument (Illumina, San Diego, CA, USA) with 75 BP single-end reads.

2.4 | Read mapping

The bcbio-nextgen pipeline v. 1.0.2 (bcbio, n.d.) was used to automate the read mapping and peak calling stages of the analysis. Raw sequence reads were mapped to the *S. cerevisiae* reference genome (SacCer3) using the Bowtie2 read mapper, version 2.2.8 (Langmead & Salzberg, 2012), using default settings. Read mapping assessment and quality checking were performed with the following tools: Samtools version 1.4 (H. Li et al., 2009), Sambamba v. 0.6.6 (Tarasov, Vilella, Cuppen, Nijman, & Prins, 2015), BEDTools v. 2.26.0 (Quinlan & Hall, 2010) and FastQC v. 0.11.5 (FastQC, n.d.).

Mapped sets of reads were down-sampled to a fixed level of 10 million reads per sample in order to collect metrics for unmapped, multi-mapping, single-mapping and duplicate reads. Unique (non-duplicate) reads that map to a single genomic location were retained from the original full sets of reads and a further down-sampling to 3 million reads was performed, in order to enter peak calling with the same coverage level for all samples.

2.5 | Peak calling

Peak calling was carried out with Bcbio, using MACS2 version 2.1.1.20160309 (Zhang et al., 2008) with the following settings: -q

0.1, $-m$ 2 50, $--bw$ 186, $-g$ 12000000. The bandwidth parameter was set to the modal value of the size of sheared DNA after sonication. Each FAIRE sample was analysed together with its corresponding input control. Overlapping peaks between each sample group (500, 100, 20 or 4 mg cell starting amounts) were identified using the *intersect* tool from the BEDTools package (Quinlan & Hall, 2010).

2.6 | IDR (irreproducible discovery rate) analysis

To perform IDR analysis (Q. Li, Brown, Huang, & Bickel, 2011), peaks called with MACS2 were analysed by the IDR protocol (IDR, n.d.). The number of significant peaks across the replicates for IDR rates of 0–30% was calculated in 0.1% increments.

2.7 | Mapping of genomic features

A summary of peak distribution over different chromosomal features was prepared using the *assignChromosomeRegions* function of the ChIPpeakAnno Bioconductor package (Zhu, 2013).

3 | RESULTS AND DISCUSSION

FAIRE-seq was performed on a series of diminishing starting amounts of yeast cells from 3-day-old colonies. The highest input amount, 500 mg cells (wet weight), was chosen as a generous amount from which a high-quality dataset was anticipated, and is approximately equivalent to the amount of cells that can be expected from 50 mL liquid cultures that have been published to date (Berchowitz et al., 2009; Connelly et al., 2014), estimated to be $\sim 1 \times 10^9$ cells. Thereafter, 5-fold decreasing amounts of cells were used to test the performance limits of the method over a broad range of inputs. The lowest amount tested here, 4 mg, thus lowers the required input cell numbers by ~ 100 -fold relative to previously published FAIRE procedures.

FAIRE was performed on three replicates at each starting cell amount, and the resulting DNA fragments, enriched for open chromatin regions, were analysed by high-throughput sequencing. For each replicate sample, an input DNA sample (i.e. cross-linked, then de-cross-linked and subjected to DNA isolation) was prepared and sequenced in parallel to control for DNA amplification artefacts that could otherwise be mistaken for FAIRE regions of open chromatin.

Sequencing generated 10–30 million reads per sample. To ensure a fair comparison of read yields between samples, all datasets were therefore down-sampled to 10 million raw reads before further processing. Read mapping revealed a similar distribution of uniquely mapping, duplicate and non-mapping reads in data from all starting cell amounts (Figure 1). Previous studies of scaled-down method performance have revealed that lower inputs can result in an abundance of un-mapped and duplicate reads (Brind'Amour et al., 2015; Gilfillan et al., 2012). However, within the ranges tested here, the input cell number had no effect on read mapping, and all samples produced highly satisfactory numbers of uniquely mapping (non-duplicate) reads.

We therefore proceeded to call peaks on the datasets using MACS2 software (Zhang et al., 2008). The corresponding input sample for each FAIRE sample was used as a control when calling peaks. Between 4200 and 6400 peaks were called in the individual samples,

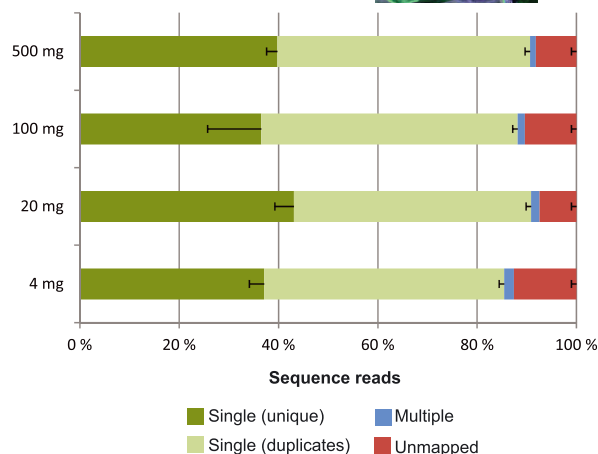


FIGURE 1 Genomic mapping of sequence reads. The proportions of unmapped reads (red), those mapping to single genomic positions (green), and those mapping to multiple locations (repeats, in blue) are shown. Reads mapped to single genomic positions are divided into those present as a unique copy and those present in two or more identical copies (duplicates). Results shown are the mean of three replicates for each method, using 10 million raw reads per replicate. Error bars show the standard deviation from the mean [Colour figure can be viewed at wileyonlinelibrary.com]

which corresponds to approximately one peak per gene in the yeast genome. Peaks were primarily found in promoter regions, consistent with expectations that FAIRE preferentially isolates regulatory DNA elements (Figure 2). Although fewest peaks were called in the 4 mg samples, there was not a consistent drop in peak number as input amounts were decreased (Table 1). There was also no indication in the data that peak calling was less reproducible as cell numbers decreased, as the standard deviation of peaks called attests. Visual

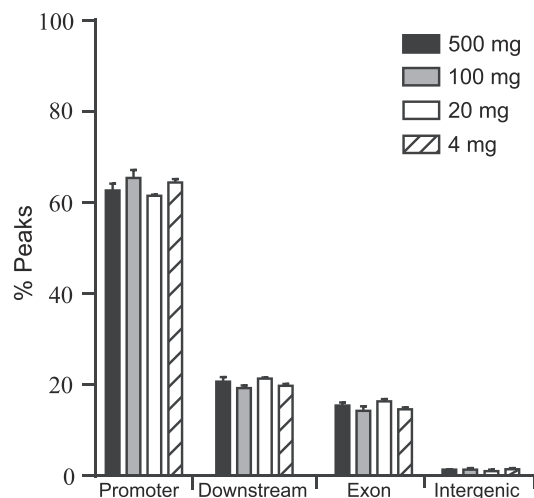
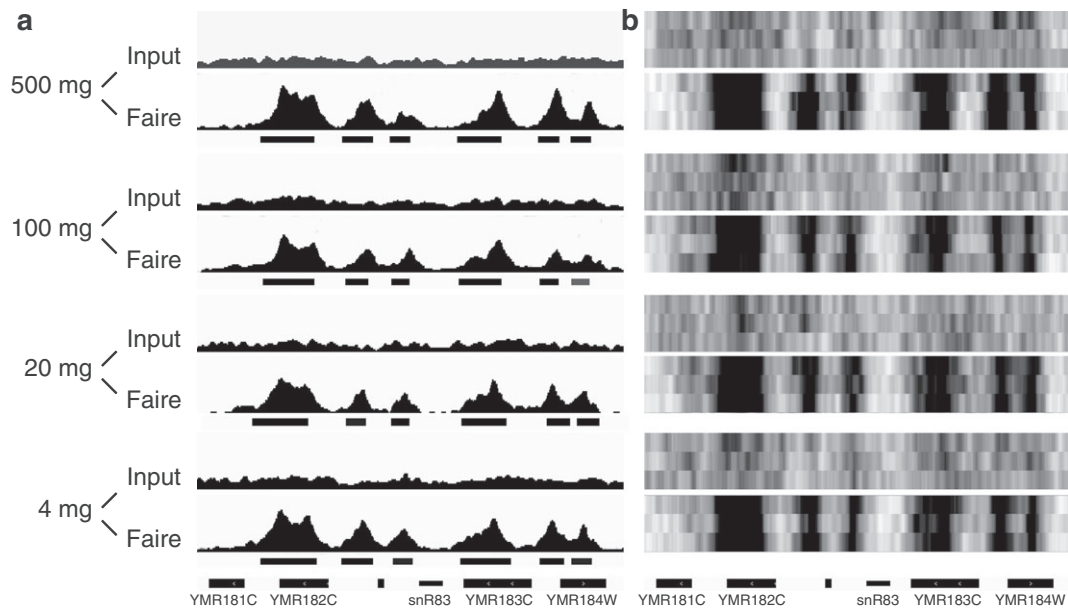


FIGURE 2 Genomic locations of FAIRE peaks. y-Axis shows percentage peaks for each given starting amount of yeast cells. Promoter regions were defined as overlapping or within 1 kb upstream of the transcription start site. Downstream regions are those within 1 kb downstream of the gene but not overlapping the 3' end of the gene. Intergenic regions were those that were > 1 kb removed from transcription start sites or gene 3' ends. No significant differences were detected between the different starting amounts (two-tailed Student's *t*-test)

TABLE 1 Peak calling, sensitivity and specificity. Peaks were called using MACS2, using 3 million uniquely mapping non-duplicate reads per sample

Input amount (mg)	Mean number of peaks called	Mean number of peaks overlapping 500 mg dataset peaks	Mean number of peaks not found in 500 mg dataset	Mean sensitivity (percentage reference peaks detected)	Mean specificity (percentage method peaks found in reference dataset)
500	5096 ± 126	4574 ± 0	522 ± 126	100	90
100	5008 ± 812	3759 ± 333	1249 ± 630	82	75
20	6292 ± 128	4284 ± 140	2008 ± 262	94	68
4	4896 ± 283	3945 ± 354	951 ± 251	86	81

Peak regions present in all three 500 mg datasets ($n = 4574$) were used as the reference dataset to which all other samples were compared to measure sensitivity and specificity. Data presented are mean ± standard deviation of three replicates.

**FIGURE 3** FAIRE-seq and input data in a 5 kb genomic region visualized in the Integrative Genomics Viewer. (a) A single example of FAIRE-seq and corresponding input sample is shown for each input amount. Peaks called by MACS2 are shown as black bars. Genes are shown at the bottom of the figure. (b) Data for all replicates shown as heat maps of the same genomic region. y-Axis scale in all cases is read depth 0–100

browsing of the data in the Integrative Genomics Viewer (Robinson et al., 2011; Figure 3) confirmed the quality of all datasets, and clearly demonstrated a good signal-to-noise ratio in FAIRE datasets and scarcity of amplification artefacts in input controls.

To assess sensitivity (false negatives) and specificity (false positives) as input amounts were decreased, we designated those peaks common to all three 500 mg datasets as a 'reference' to which the remaining datasets could be compared (Table 1). All three lower input amounts faithfully detected 82–94% of the peaks present in the 500 mg reference dataset. Despite having the lowest number of overall peaks called, the 4 mg dataset did not detect fewer reference peaks. Specificity (a low value indicates the presence of additional peaks not present in the reference dataset) was also high in all samples, and once again was not lowest in the 4 mg dataset.

To assess reproducibility, we examined the differences in peak calls within each input amount (Figure 4a) and between the input amounts (Figure 4b). The least variation was seen between replicates of the 500 mg input samples (Figure 4a). However, variation did not consistently increase with decreasing input amount, and individual replicate samples prepared from 4 mg input appeared superior to those prepared from 100 mg input. When limiting the comparison to

only peaks present in all three replicates from each input amount (Figure 4b), the majority of peaks were detected at all input amounts. To more objectively measure reproducibility, we performed IDR analysis (Q. Li et al., 2011) of the peak calling data. The number of significant peaks across the replicates for different IDR rates (0–30% in 0.1% increments) was calculated (Figure 4c). The results show that the 500 mg dataset has the highest reproducibility, but all three other input amounts also had good reproducibility, and there was no clear difference in performance between the lower input amounts.

The conclusion of the above analyses is that no consistent drop in performance, measured in terms of read mapping, sensitivity, specificity or reproducibility, could be attributed to decreasing input cell numbers over a ~100-fold range, and the method performs well with as little as 4 mg cells input. Indeed, the FAIRE assay appears surprisingly robust to input cell numbers over the range tested here, although it undeniably worked best at the 500 mg input amount. Attempts to further increase the reproducibility of the data by adding yet more replicates did not improve the outcome (data not shown), suggesting that experimental variation has more effect than input cell number at these amounts.

To the best of our knowledge, the method presented here represents an ~ 100-fold decrease in the input cell amount compared with

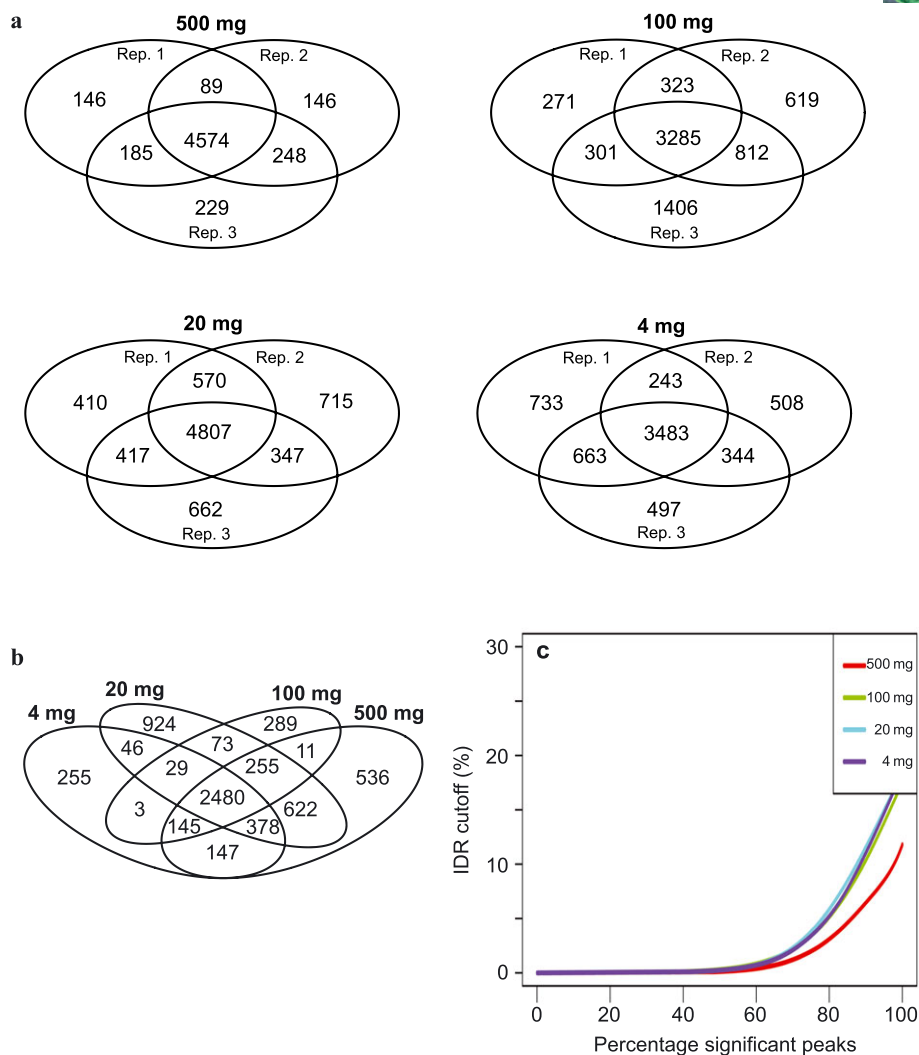


FIGURE 4 Reproducibility of peak calls. (a) Venn diagrams showing the overlapping peak calls in the triplicate samples within each input amount. (b) Venn diagram of inter-group peak overlaps, using only peaks common to all three replicates at each input amount. (c) Reproducibility measured by the irreproducible discovery rate (IDR) at different numbers of selected peaks, plotted at various IDR cutoffs. High reproducibility produces a curve with a late transition to high IDR values

earlier yeast FAIRE methods (Berchowitz et al., 2009; Connelly et al., 2014). In addition, it is the first protocol to demonstrate FAIRE on cells taken directly from colonies cultured on agar plates, rather than from liquid culture. We underline that the performance of the method has not been tested on logarithmically growing liquid cultured cells, which have typically been used for chromatin accessibility studies to date. Our goal in developing this method was to enable comparison of cells from aged and biofilm-model colony-forming yeasts (manuscript in preparation). However, the advantages of a procedure that caters to low-input amounts, does not require liquid culture, and that does not rely on incubation to generate spheroplasts extend beyond this goal: with a requirement of only 4 mg cells, this procedure should be amenable to samples of wild yeast collected from environmental samples. Furthermore, epigenetic states are sensitive to changes in environment and culture conditions, so the ability to analyse chromatin accessibility in yeast samples fixed in formaldehyde immediately after isolation is particularly important for the study of epigenetic states. Finally, as there was no dramatic deterioration in data quality

observed with decreasing cell numbers in this study, it appears we have not approached the lower limits of the method, so further decreases in input cell number may be possible.

3.1 | Availability of data

The raw data presented in this study have been archived in the NCBI's Gene Expression Omnibus repository under accession number GSE104124 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104124>).

FUNDING

The research leading to these results has received funding from the Norway Grants Norwegian Financial Mechanism 2009–2014 project no. 7F14083 under contract no. MSMT-28477/2014. Part of the research was performed in BIOCEV supported by CZ.1.05/1.1.00/02.0109 BIOCEV provided by ERDF and MEYS.

ACKNOWLEDGEMENTS

We would like to thank Arvind Y.M. Sundaram and Pål Marius Bjørnstad for advice and assistance with bioinformatic analysis. High-throughput sequencing was performed by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and Oslo University Hospital and supported by the 'Functional Genomics' and 'Infrastructure' programs of the Research Council of Norway and the Southeastern Regional Health Authorities. Computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (project no. LM2015042) and CERIT- Scientific Cloud (project no. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

ORCID

Gregor D. Gilfillan  <http://orcid.org/0000-0001-8116-0134>

REFERENCES

- Auerbach, R. K., Euskirchen, G., Rozowsky, J., Lamarre-Vincent, N., Moqtaderi, Z., Lefrancois, P., ... Snyder, M. (2009). Mapping accessible chromatin regions using Sono-Seq. *Proceedings of the National Academy of Sciences USA*, *106*(35), 14926–14931. <https://doi.org/10.1073/pnas.0905443106>
- bcbio. (n.d.)<https://github.com/chapmanb/bcbio-nextgen> (accessed May 2017)
- Berchowitz, L. E., Hanlon, S. E., Lieb, J. D., & Copenhaver, G. P. (2009). A positive but complex association between meiotic double-strand break hotspots and open chromatin in *Saccharomyces cerevisiae*. *Genome Research*, *19*(12), 2245–2257. <https://doi.org/10.1101/gr.096297.109>
- Boyle, A. P., Davis, S., Shulha, H. P., Meltzer, P., Margulies, E. H., Weng, Z., ... Crawford, G. E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell*, *132*(2), 311–322. <https://doi.org/10.1016/j.cell.2007.12.014>
- Brind'Amour, J., Liu, S., Hudson, M., Chen, C., Karimi, M. M., & Lorincz, M. C. (2015). An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nature Communications*, *6*, 6033. <https://doi.org/10.1038/ncomms7033>
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, *10*(12), 1213–1218. <https://doi.org/10.1038/nmeth.2688>
- Connelly, C. F., Wakefield, J., & Akey, J. M. (2014). Evolution and genetic architecture of chromatin accessibility and function in yeast. *PLoS Genetics*, *10*(7), e1004427. <https://doi.org/10.1371/journal.pgen.1004427>
- FastQC. (n.d.)<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed January 2017
- Gaulton, K. J., Nammo, T., Pasquali, L., Simon, J. M., Giresi, P. G., Fogarty, M. P., ... Ferrer, J. (2010). A map of open chromatin in human pancreatic islets. *Nature Genetics*, *42*(3), 255–259. <https://doi.org/10.1038/ng.530>
- Gilfillan, G. D., Hughes, T., Sheng, Y., Hjorthaug, H. S., Straub, T., Gervin, K., ... Lyle, R. (2012). Limitations and possibilities of low cell number ChIP-seq. *BMC Genomics*, *13*, 645. <https://doi.org/10.1186/1471-2164-13-645>
- Giresi, P. G., Kim, J., McDaniel, R. M., Iyer, V. R., & Lieb, J. D. (2007). FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Research*, *17*(6), 877–885. <https://doi.org/10.1101/gr.5533506>
- Hogan, G. J., Lee, C. K., & Lieb, J. D. (2006). Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. *PLoS Genetics*, *2*(9), e158. <https://doi.org/10.1371/journal.pgen.0020158>
- IDR. (n.d.)<https://sites.google.com/site/anshulkundaje/projects/idr> (accessed May 2017).
- Klis, F. M., Mol, P., Hellingwerf, K., & Brul, S. (2002). Dynamics of cell wall structure in *Saccharomyces cerevisiae*. *FEMS Microbiology Review*, *26*(3), 239–256.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, Q., Brown, J., Huang, H., & Bickel, P. (2011). Measuring reproducibility of high-throughput experiments. *Annals of Applied Statistics*, *5*(3), 28. <https://doi.org/10.1214/11-AOAS466>
- Nagy, P. L., Cleary, M. L., Brown, P. O., & Lieb, J. D. (2003). Genomewide demarcation of RNA polymerase II transcription units revealed by physical fractionation of chromatin. *Proceedings of the National Academy of Sciences USA*, *100*(11), 6364–6369. <https://doi.org/10.1073/pnas.1131966100>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Robinson, J. T., Thorvaldsdottir, H., Wincker, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24–26. <https://doi.org/10.1038/nbt.1754>
- Schep, A. N., Buenrostro, J. D., Denny, S. K., Schwartz, K., Sherlock, G., & Greenleaf, W. J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Research*, *25*(11), 1757–1770. <https://doi.org/10.1101/gr.192294.115>
- Schones, D. E., Cui, K., Cuddapah, S., Roh, T. Y., Barski, A., Wang, Z., ... Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell*, *132*(5), 887–898. <https://doi.org/10.1016/j.cell.2008.02.022>
- Simon, J. M., Giresi, P. G., Davis, I. J., & Lieb, J. D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nature Protocols*, *7*(2), 256–267. <https://doi.org/10.1038/nprot.2011.444>
- Simon, J. M., Giresi, P. G., Davis, I. J., & Lieb, J. D. (2013). A detailed protocol for formaldehyde-assisted isolation of regulatory elements (FAIRE). *Current Protocols in Molecular Biology*. Chapter 21, Units 21–26. <https://doi.org/10.1002/0471142727.mb2126s102>
- Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: Fast processing of NGS alignment formats. *Bioinformatics*, *31*(12), 2032–2034. <https://doi.org/10.1093/bioinformatics/btv098>
- Tsompana, M., & Buck, M. J. (2014). Chromatin accessibility: A window into the genome. *Epigenetics & Chromatin*, *7*(1), 33. <https://doi.org/10.1186/1756-8935-7-33>
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., ... Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, *9*(9), R137. <https://doi.org/10.1186/gb-2008-9-9-r137>
- Zhu, L. J. (2013). Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods in Molecular Biology*, *1067*, 105–124. https://doi.org/10.1007/978-1-62703-607-8_8

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Segorbe D, Wilkinson D, Mizeranschi A, et al. An optimized FAIRE procedure for low cell numbers in yeast. *Yeast*. 2018;35:507–512. <https://doi.org/10.1002/yea.3316>